

# CONJUGATION IN A GROUP

KEITH CONRAD

## 1. INTRODUCTION

A reflection across one line in the plane is, geometrically, just like a reflection across every other line. That is, while reflections across two different lines in the plane are not strictly the same, they have the same *type* of effect. Similarly, two different transpositions in  $S_n$  are not the same permutation but have the same *type* of effect: swap two elements and leave everything else unchanged. The concept that makes the notion of “different, but same type of effect” precise is called conjugacy.

In a group  $G$ , two elements  $g$  and  $h$  are called *conjugate* when

$$h = xgx^{-1}$$

for some  $x \in G$ . This relation is symmetric, since  $g = yhy^{-1}$  with  $y = x^{-1}$ . When  $h = xgx^{-1}$ , we say  $x$  conjugates  $g$  to  $h$ . (Warning: when some people say “ $x$  conjugates  $g$  to  $h$ ” they might mean  $h = x^{-1}gx$  instead of  $h = xgx^{-1}$ .)

**Example 1.1.** The table below lists all  $\sigma(12)\sigma^{-1}$  for  $\sigma \in S_3$ .

$\sigma$	(1)	(12)	(13)	(23)	(123)	(132)
$\sigma(12)\sigma^{-1}$	(12)	(12)	(23)	(13)	(23)	(13)

The conjugates of (12) are in the second row: (12), (13), and (23). Notice the redundancy in the table: each conjugate of (12) arises in two ways.

We will see in Theorem 5.4 that in  $S_n$  all transpositions are conjugate to each other. In Appendix A is a proof that the reflections across each of two lines in the plane are conjugate to each other in the group of all isometries of the plane.

It is useful to collect conjugate elements in a group together, and these are called conjugacy classes. Examples of them are in Section 2. Some theorems about conjugate elements are proved in Section 3. We’ll look at conjugate elements of  $D_n$  in Section 4 and conjugate permutations in  $S_n$  and  $A_n$  in Section 5. In Section 6 we will introduce some subgroups that are related to conjugacy and use them to prove some theorems about finite  $p$ -groups, such as a classification of groups of order  $p^2$  and the existence of a normal (!) subgroup of every order dividing the order of a  $p$ -group.

## 2. CONJUGACY CLASSES: DEFINITION AND EXAMPLES

For an element  $g$  of a group  $G$ , its *conjugacy class* is the set of elements conjugate to it:

$$\{xgx^{-1} : x \in G\}.$$

**Example 2.1.** If  $G$  is abelian then  $xgx^{-1} = g$  for all  $x, g \in G$ : every  $g$  is its own conjugacy class. This characterizes abelian groups: to say each  $g \in G$  is its own conjugacy class means  $xgx^{-1} = g$  for all  $x$  and  $g$  in  $G$ , which says  $xg = gx$  for all  $x$  and  $g$ , so  $G$  is abelian.

**Example 2.2.** The conjugacy class of  $(12)$  in  $S_3$  is  $\{(12), (13), (23)\}$ , as we saw in Example 1.1. Similarly, the reader can check the conjugacy class of  $(123)$  is  $\{(123), (132)\}$ . The conjugacy class of  $(1)$  is just  $\{(1)\}$ . So  $S_3$  has three conjugacy classes:

$$\{(1)\}, \{(12), (13), (23)\}, \{(123), (132)\}.$$

**Example 2.3.** In  $D_4 = \langle r, s \rangle$ , there are five conjugacy classes:

$$\{1\}, \{r^2\}, \{s, r^2s\}, \{r, r^3\}, \{rs, r^3s\}.$$

The members of a conjugacy class of  $D_4$  are different but have the same type of effect on a square:  $r$  and  $r^3$  are a 90 degree rotation in some direction,  $s$  and  $r^2s$  are a reflection across a diagonal, and  $rs$  and  $r^3s$  are a reflection across an edge bisector.

**Example 2.4.** There are five conjugacy classes in  $Q_8$ :

$$\{1\}, \{-1\}, \{i, -i\}, \{j, -j\}, \{k, -k\}.$$

**Example 2.5.** There are four conjugacy classes in  $A_4$ :

$$\begin{aligned} &\{(1)\}, \{(12)(34), (13)(24), (14)(23)\}, \\ &\{(123), (243), (134), (142)\}, \{(132), (234), (143), (124)\}. \end{aligned}$$

Notice the 3-cycles  $(123)$  and  $(132)$  are *not* conjugate in  $A_4$ . All 3-cycles in  $A_4$  are conjugate in the larger group  $S_4$ , *e.g.*,  $(132) = (23)(123)(23)^{-1}$  and the conjugating permutation  $(23)$  is not in  $A_4$ .

In these examples, different conjugacy classes in a group are *disjoint*: they don't overlap at all. This will be proved in general in Section 3. Also, the sizes of different conjugacy classes are not all the same, but these sizes all divide the size of the group. We will see in Section 6 why this is true.

The idea of conjugation can be applied not just to elements, but to subgroups. If  $H \subset G$  is a subgroup and  $g \in G$ , the set

$$gHg^{-1} = \{ghg^{-1} : h \in H\}$$

is a subgroup of  $G$ , called naturally enough a *conjugate subgroup* to  $H$ . It's a subgroup since it contains the identity ( $e = geg^{-1}$ ) and is closed under multiplication and inversion:  $(ghg^{-1})(gh'g^{-1}) = g(hh')g^{-1}$  and  $(ghg^{-1})^{-1} = gh^{-1}g^{-1}$ . Unlike different conjugacy classes, different conjugate subgroups are not disjoint: they all contain the identity.

**Example 2.6.** While  $D_4$  has 5 conjugacy classes of elements (Example 2.3), it has 8 conjugacy classes of subgroups. In total there are 10 subgroups of  $D_4$ :

$$\langle 1 \rangle = \{1\}, \quad \langle s \rangle = \{1, s\}, \quad \langle rs \rangle = \{1, rs\}, \quad \langle r^2s \rangle = \{1, r^2s\}, \quad \langle r^3s \rangle = \{1, r^3s\},$$

$$\langle r \rangle = \{1, r, r^2, r^3\}, \quad \langle r^2 \rangle = \{1, r^2\}, \quad \langle r^2, s \rangle = \{1, r^2, s, r^2s\}, \quad \langle r^2, rs \rangle = \{1, r^2, rs, r^3s\}, \quad D_4.$$

In this list the subgroups  $\langle s \rangle$  and  $\langle r^2s \rangle$  are conjugate, as are  $\langle rs \rangle$  and  $\langle r^3s \rangle$ : check  $r\langle s \rangle r^{-1} = \langle r^2s \rangle$  and  $r\langle rs \rangle r^{-1} = \langle r^3s \rangle$ . The other six subgroups of  $D_4$  are conjugate only to themselves.

We will not discuss conjugate subgroups much, but the concept is important. For instance, a subgroup is conjugate only to itself precisely when it is a normal subgroup.

3. SOME BASIC PROPERTIES OF CONJUGACY CLASSES

**Lemma 3.1.** *In a group,  $(xgx^{-1})^n = xg^n x^{-1}$  for all positive integers  $n$ .*

*Proof.* This is left to the reader as an exercise using induction. The equation is in fact true for all  $n \in \mathbf{Z}$ .  $\square$

**Theorem 3.2.** *All the elements of a conjugacy class have the same order.*

*Proof.* This is saying  $g$  and  $xgx^{-1}$  have the same order. By Lemma 3.1,  $(xgx^{-1})^n = xg^n x^{-1}$  for all  $n \in \mathbf{Z}^+$ , so if  $g^n = 1$  then  $(xgx^{-1})^n = xg^n x^{-1} = xx^{-1} = e$ , and if  $(xgx^{-1})^n = 1$  then  $xg^n x^{-1} = e$ , so  $g^n = x^{-1}x = e$ . Thus  $(xgx^{-1})^n = 1$  if and only if  $g^n = 1$ , so  $g$  and  $xgx^{-1}$  have the same order.  $\square$

The converse to Theorem 3.2 is false: elements of the same order in a group need not be conjugate in that group. This is clear in abelian groups, where different elements are never conjugate but could have the same order. Looking at the nonabelian examples in Section 2, in  $D_4$  there are five elements of order two spread across 3 conjugacy classes. Similarly, there are non-conjugate elements of equal order in  $Q_8$  and  $A_4$ . But in  $S_3$ , elements of equal order in  $S_3$  are conjugate. Amazingly, this is the largest example of a finite group where that property holds: up to isomorphism, the only nontrivial finite groups where all elements of equal order are conjugate are  $\mathbf{Z}/(2)$  and  $S_3$ . A proof is given in [3] and [6], and depends on the classification of finite simple groups. A conjugacy problem about  $S_3$  that remains open, as far as I know, is the conjecture that  $S_3$  is the only nontrivial finite group (up to isomorphism) in which different conjugacy classes all have different sizes.

**Corollary 3.3.** *If  $H$  is a cyclic subgroup of  $G$  then every subgroup conjugate to  $H$  is cyclic.*

*Proof.* Writing  $H = \langle y \rangle = \langle y^n : n \in \mathbf{Z} \rangle$ ,

$$gHg^{-1} = \{gy^n g^{-1} : n \in \mathbf{Z}\} = \{(gyg^{-1})^n : n \in \mathbf{Z}\} = \langle gyg^{-1} \rangle,$$

so  $gHg^{-1}$  is cyclic with a generator being a conjugate (by  $g$ ) of a generator of  $H$ .  $\square$

Let's verify the observation in Section 2 that different conjugacy classes are disjoint.

**Theorem 3.4.** *Let  $G$  be a group and  $g, h \in G$ . If the conjugacy classes of  $g$  and  $h$  overlap, then the conjugacy classes are equal.*

*Proof.* We need to show every element conjugate to  $g$  is also conjugate to  $h$ , and *vice versa*. Since the conjugacy classes overlap, we have  $xgx^{-1} = yhy^{-1}$  for some  $x$  and  $y$  in the group. Therefore

$$g = x^{-1}yhy^{-1}x = (x^{-1}y)h(x^{-1}y)^{-1},$$

so  $g$  is conjugate to  $h$ . Each element conjugate to  $g$  is  $zgz^{-1}$  for some  $z \in G$ , and

$$zgz^{-1} = z(x^{-1}y)h(x^{-1}y)^{-1}z^{-1} = (zx^{-1}y)h(zx^{-1}y)^{-1},$$

which shows each element of  $G$  that is conjugate to  $g$  is also conjugate to  $h$ . To go the other way, from  $xgx^{-1} = yhy^{-1}$  write  $h = (y^{-1}x)g(y^{-1}x)^{-1}$  and carry out a similar calculation.  $\square$

Theorem 3.4 says each element of a group belongs to just one conjugacy class. We call an element of a conjugacy class a *representative* of that class.

A conjugacy class consists of all  $xgx^{-1}$  for fixed  $g$  and varying  $x$ . Instead we can look at all  $xgx^{-1}$  for fixed  $x$  and varying  $g$ . That is, instead of looking at all the elements

conjugate to  $g$  we look at all the ways  $x$  can conjugate the elements of the group. This “conjugate-by- $x$ ” function is denoted  $\gamma_x: G \rightarrow G$ , so  $\gamma_x(g) = xgx^{-1}$ .

**Theorem 3.5.** *Each conjugation function  $\gamma_x: G \rightarrow G$  is an automorphism of  $G$ .*

*Proof.* For all  $g$  and  $h$  in  $G$ ,

$$\gamma_x(g)\gamma_x(h) = xgx^{-1}xhx^{-1} = xghx^{-1} = \gamma_x(gh),$$

so  $\gamma_x$  is a homomorphism. Since  $h = xgx^{-1}$  if and only if  $g = x^{-1}hx$ , the function  $\gamma_x$  has inverse  $\gamma_{x^{-1}}$ , so  $\gamma_x$  is an automorphism of  $G$ .  $\square$

Theorem 3.5 explains why conjugate elements in a group  $G$  are “the same except for the point of view”: they are linked by an automorphism of  $G$ , namely some map  $\gamma_x$ . This means an element of  $G$  and its conjugates in  $G$  have the same group-theoretic properties, such as: having the same order, being an  $m$ -th power, being in the center, and being a commutator. Likewise, a subgroup  $H$  and its conjugates  $gHg^{-1}$  have the same group-theoretic properties.

Automorphisms of  $G$  having the form  $\gamma_x$  are called *inner automorphisms*. They are the only automorphisms that can be written down without knowing extra information about  $G$  (such as being told  $G$  is abelian or that  $G$  is a particular matrix group). For some  $G$  every automorphism of  $G$  is an inner automorphism. This is true for the groups  $S_n$  when  $n \neq 6$  (that’s right:  $S_6$  is the only symmetric group with an automorphism that isn’t conjugation by a permutation). The groups  $\mathrm{GL}_n(\mathbf{R})$  when  $n \geq 2$  have extra automorphisms: since  $(AB)^\top = B^\top A^\top$  and  $(AB)^{-1} = B^{-1}A^{-1}$ , the function  $f(A) = (A^\top)^{-1}$  on  $\mathrm{GL}_n(\mathbf{R})$  is an automorphism and it is not inner.

Here is a simple result where inner automorphisms tell us something about all automorphisms of a group.

**Theorem 3.6.** *If  $G$  is a group with trivial center, then the group  $\mathrm{Aut}(G)$  also has trivial center.*

*Proof.* Let  $\varphi \in \mathrm{Aut}(G)$  and assume  $\varphi$  commutes with all other automorphisms. We will see what it means for  $\varphi$  to commute with an inner automorphism  $\gamma_x$ . For  $g \in G$ ,

$$(\varphi \circ \gamma_x)(g) = \varphi(\gamma_x(g)) = \varphi(xgx^{-1}) = \varphi(x)\varphi(g)\varphi(x)^{-1}$$

and

$$(\gamma_x \circ \varphi)(g) = \gamma_x(\varphi(g)) = x\varphi(g)x^{-1},$$

so having  $\varphi$  and  $\gamma_x$  commute means, for all  $g \in G$ , that

$$\varphi(x)\varphi(g)\varphi(x)^{-1} = x\varphi(g)x^{-1} \iff x^{-1}\varphi(x)\varphi(g) = \varphi(g)x^{-1}\varphi(x),$$

so  $x^{-1}\varphi(x)$  commutes with every value of  $\varphi$ . Since  $\varphi$  is onto,  $x^{-1}\varphi(x) \in Z(G)$ . The center of  $G$  is trivial, so  $\varphi(x) = x$ . This holds for all  $x \in G$ , so  $\varphi$  is the identity automorphism. We have proved the center of  $\mathrm{Aut}(G)$  is trivial.  $\square$

#### 4. CONJUGACY CLASSES IN $D_n$

In the group  $D_n$  we will show rotations are conjugate only to their inverses and reflections are either all conjugate or fall into two conjugacy classes.

**Theorem 4.1.** *The conjugacy classes in  $D_n$  are as follows.*

- (1) *If  $n$  is odd,*
  - *the identity element:  $\{1\}$ ,*

- $(n - 1)/2$  conjugacy classes of size 2:  $\{r^{\pm 1}\}, \{r^{\pm 2}\}, \dots, \{r^{\pm(n-1)/2}\},$
  - all the reflections:  $\{r^i s : 0 \leq i \leq n - 1\}.$
- (2) If  $n$  is even,
- two conjugacy classes of size 1:  $\{1\}, \{r^{\frac{n}{2}}\},$
  - $n/2 - 1$  conjugacy classes of size 2:  $\{r^{\pm 1}\}, \{r^{\pm 2}\}, \dots, \{r^{\pm(\frac{n}{2}-1)}\},$
  - the reflections fall into two conjugacy classes:  $\{r^{2i} s : 0 \leq i \leq \frac{n}{2} - 1\}$  and  $\{r^{2i+1} s : 0 \leq i \leq \frac{n}{2} - 1\}.$

*Proof.* Every element of  $D_n$  is  $r^i$  or  $r^i s$  for some integer  $i$ . Therefore to find the conjugacy class of an element  $g$  we will compute  $r^i g r^{-i}$  and  $(r^i s) g (r^i s)^{-1}$ .

The formulas

$$r^i r^j r^{-i} = r^j, \quad (r^i s) r^j (r^i s)^{-1} = r^{-j}$$

as  $i$  varies show the only conjugates of  $r^j$  in  $D_n$  are  $r^j$  and  $r^{-j}$ . Explicitly, the basic formula  $s r^j s^{-1} = r^{-j}$  shows us  $r^j$  and  $r^{-j}$  are conjugate; we need the more general calculation to be sure there is nothing further that  $r^j$  is conjugate to.

To find the conjugacy class of  $s$ , we compute

$$r^i s r^{-i} = r^{2i} s, \quad (r^i s) s (r^i s)^{-1} = r^{2i} s.$$

As  $i$  varies,  $r^{2i} s$  runs through the reflections in which  $r$  occurs with an exponent divisible by 2. If  $n$  is odd then every integer modulo  $n$  is a multiple of 2 (since 2 is invertible mod  $n$  we can solve  $k \equiv 2i \pmod{n}$  for  $i$  no matter what  $k$  is). Therefore when  $n$  is odd

$$\{r^{2i} s : i \in \mathbf{Z}\} = \{r^k s : k \in \mathbf{Z}\},$$

so every reflection in  $D_n$  is conjugate to  $s$ . When  $n$  is even, however, we only get half the reflections as conjugates of  $s$ . The other half are conjugate to  $rs$ :

$$r^i (rs) r^{-i} = r^{2i+1} s, \quad (r^i s) (rs) (r^i s)^{-1} = r^{2i-1} s.$$

As  $i$  varies, this gives us  $\{rs, r^3 s, \dots, r^{n-1} s\}$ . □

That reflections in  $D_n$  form either one or two conjugacy classes, depending on the parity of  $n$ , corresponds to a geometric feature of reflections: for odd  $n$  all reflections in  $D_n$  look the same (Figure 1) – reflecting across a line connecting a vertex and the midpoint on the opposite side – but for even  $n$  the reflections in  $D_n$  fall into two types – the  $r^{\text{even}} s$  reflect across a line through pairs of opposite vertices and the  $r^{\text{odd}} s$  reflect across a line through midpoints of opposite sides (Figure 2).

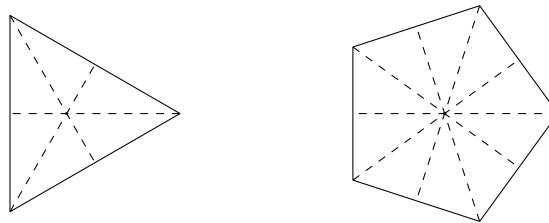
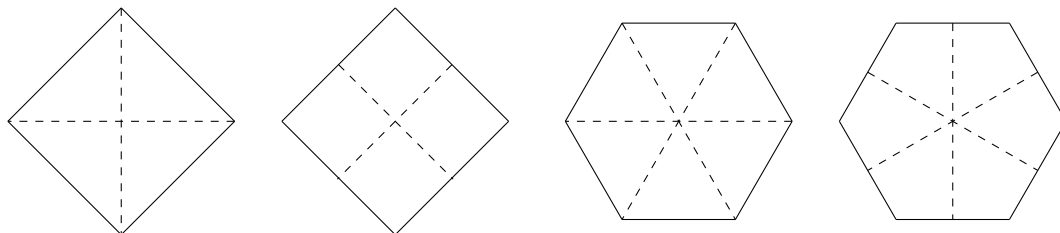


FIGURE 1. Lines of Reflection for  $n = 3$  and  $n = 5$ .

FIGURE 2. Lines of Reflection for  $n = 4$  and  $n = 6$ .5. CONJUGACY CLASSES IN  $S_n$  AND  $A_n$ 

The following tables list a representative from each conjugacy class in  $S_n$  and  $A_n$  for  $3 \leq n \leq 6$ , along with the size of the conjugacy classes. Conjugacy classes disjointly cover a group, by Theorem 3.4, so the conjugacy class sizes add up to  $n!$  for  $S_n$  and  $n!/2$  for  $A_n$ .

	$S_3$			$A_3$		
Rep.	(1)	(123)	(12)	(1)	(123)	(132)
Size	1	2	3	1	1	1

	$S_4$					$A_4$			
Rep.	(1)	(12)(34)	(12)	(1234)	(123)	(1)	(12)(34)	(123)	(132)
Size	1	3	6	6	8	1	3	4	4

	$S_5$						
Rep.	(1)	(12)	(12)(34)	(123)	(12)(345)	(12345)	(1234)
Size	1	10	15	20	20	24	30

	$A_5$				
Rep.	(1)	(12345)	(21345)	(12)(34)	(123)
Size	1	12	12	15	20

	$S_6$					
Rep.	(1)	(12)	(12)(34)(56)	(123)	(123)(456)	(12)(34)
Size	1	15	15	40	40	45
Rep.	(1234)	(12)(3456)	(123456)	(12)(345)	(12345)	
Size	90	90	120	120	144	

	$A_6$						
Rep.	(1)	(123)	(123)(456)	(12)(34)	(12345)	(23456)	(1234)(56)
Size	1	40	40	45	72	72	90

Notice elements of  $A_n$  can be conjugate in  $S_n$  while *not* being conjugate in  $A_n$ , such as (123) and (132) for  $n = 3$  and  $n = 4$ . (See Example 2.5.) The permutations in  $S_3$  and  $S_4$  that conjugate (123) to (132) are not even, so (123) and (132) are not conjugate in  $A_3$  or  $A_4$ . They are conjugate in  $A_5$ :  $(132) = \sigma(123)\sigma^{-1}$  for  $\sigma = (23)(45)$ .

As a first step in describing conjugacy classes in  $S_n$ , let's find the conjugates of a  $k$ -cycle.

**Theorem 5.1.** For each cycle  $(i_1 i_2 \dots i_k)$  in  $S_n$  and each  $\sigma \in S_n$ ,

$$\sigma(i_1 i_2 \dots i_k) \sigma^{-1} = (\sigma(i_1) \sigma(i_2) \dots \sigma(i_k)).$$

Before proving this formula, let's see in two examples how it works.

**Example 5.2.** In  $S_5$ , let  $\sigma = (13)(254)$ . Then

$$\sigma(1432) \sigma^{-1} = (13)(254)(1432)(245)(13) = (1532)$$

while  $(\sigma(1)\sigma(4)\sigma(3)\sigma(2)) = (3215)$  since  $\sigma(1) = 3$ ,  $\sigma(4) = 2$ ,  $\sigma(3) = 1$ , and  $\sigma(2) = 5$ . Clearly  $(1532) = (3215)$ .

**Example 5.3.** In  $S_7$ , let  $\sigma = (13)(265)$ . Then

$$\sigma(73521) \sigma^{-1} = (13)(265)(73521)(256)(13) = (12637)$$

and  $(\sigma(7)\sigma(3)\sigma(5)\sigma(2)\sigma(1)) = (71263) = (12637)$ .

Now we prove Theorem 5.1.

*Proof.* Let  $\pi = \sigma(i_1 i_2 \dots i_k) \sigma^{-1}$ . We want to show  $\pi$  is the cyclic permutation of the numbers  $\sigma(i_1), \sigma(i_2), \dots, \sigma(i_k)$ . That means two things:

- Show  $\pi$  sends  $\sigma(i_1)$  to  $\sigma(i_2)$ ,  $\sigma(i_2)$  to  $\sigma(i_3)$ ,  $\dots$ , and finally  $\sigma(i_k)$  to  $\sigma(i_1)$ .
- Show  $\pi$  does not move a number other than  $\sigma(i_1), \dots, \sigma(i_k)$ .

The second step is essential. Just knowing a permutation cyclically permutes certain numbers does not mean it *is* the cycle built from those numbers, since it could move other numbers we haven't looked at yet. (For instance, if  $\pi(1) = 2$  and  $\pi(2) = 1$ ,  $\pi$  need not be  $(12)$ . The permutation  $(12)(345)$  also has that behavior.)

What does  $\pi$  do to  $\sigma(i_1)$ ? The effect is

$$\pi(\sigma(i_1)) = (\sigma(i_1 i_2 \dots i_k) \sigma^{-1})(\sigma(i_1)) = ((\sigma(i_1 i_2 \dots i_k) \sigma^{-1} \sigma)(i_1) = \sigma(i_1 i_2 \dots i_k)(i_1) = \sigma(i_2).$$

(The " $(i_1)$ " at the ends is not a 1-cycle, but denotes the point where a permutation is being evaluated.) Similarly,  $\pi(\sigma(i_2)) = \sigma(i_1 i_2 \dots i_k)(i_2) = \sigma(i_3)$ , and so on up to  $\pi(\sigma(i_k)) = \sigma(i_1 i_2 \dots i_k)(i_k) = \sigma(i_1)$ .

Now pick a number  $a$  that is not among  $\sigma(i_1), \dots, \sigma(i_k)$ . We want to show  $\pi(a) = a$ . That means we want to show  $\sigma(i_1 i_2 \dots i_k) \sigma^{-1}(a) = a$ . Since  $a \neq \sigma(i_j)$  for  $j = 1, \dots, k$ , also  $\sigma^{-1}(a)$  is not  $i_j$  for  $j = 1, \dots, k$ . Therefore the cycle  $(i_1 i_2 \dots i_k)$  does not move  $\sigma^{-1}(a)$ , so its effect on  $\sigma^{-1}(a)$  is to keep it as  $\sigma^{-1}(a)$ . Hence

$$\pi(a) = (\sigma(i_1 i_2 \dots i_k) \sigma^{-1})(a) = \sigma(i_1 i_2 \dots i_k)(\sigma^{-1}(a)) = \sigma(\sigma^{-1}(a)) = a. \quad \square$$

We now know that every conjugate of a cycle is also a cycle of the same length. Is the converse true, *i.e.*, if two cycles have the same length are they conjugate?

**Theorem 5.4.** All cycles of the same length in  $S_n$  are conjugate.

*Proof.* Pick two  $k$ -cycles, say

$$(a_1 a_2 \dots a_k), \quad (b_1 b_2 \dots b_k).$$

Choose  $\sigma \in S_n$  so that  $\sigma(a_1) = b_1, \dots, \sigma(a_k) = b_k$ , and let  $\sigma$  be an arbitrary bijection from the complement of  $\{a_1, \dots, a_k\}$  to the complement of  $\{b_1, \dots, b_k\}$ . Then, using Theorem 5.1, we see conjugation by  $\sigma$  carries the first  $k$ -cycle to the second.  $\square$

For instance, the transpositions (2-cycles) in  $S_n$  form a single conjugacy class, as we saw for  $S_3$  in the introduction.

Now we consider the conjugacy class of an arbitrary permutation in  $S_n$ , not necessarily a cycle. It will be convenient to introduce some terminology. Writing a permutation as a product of disjoint cycles, arrange the lengths of those cycles in increasing order, including 1-cycles if there are fixed points. These lengths are called the *cycle type* of the permutation.<sup>1</sup> For instance, in  $S_7$  the permutation (12)(34)(567) is said to have cycle type (2, 2, 3). When discussing the cycle type of a permutation, we include fixed points as 1-cycles. For instance, (12)(35) in  $S_5$  is (4)(12)(35) and has cycle type (1, 2, 2). If we view (12)(35) in  $S_6$  then it is (4)(6)(12)(35) and has cycle type (1, 1, 2, 2).

The cycle type of a permutation in  $S_n$  is just a set of positive integers that add up to  $n$ , which is called a *partition* of  $n$ . There are 7 partitions of 5:

$$5, 1 + 4, 2 + 3, 1 + 1 + 3, 1 + 2 + 2, 1 + 1 + 1 + 2, 1 + 1 + 1 + 1 + 1.$$

Thus, the permutations of  $S_5$  have 7 cycle types. Knowing the cycle type of a permutation tells us its disjoint cycle structure except for how the particular numbers fall into the cycles. For instance, a permutation in  $S_5$  with cycle type (1, 2, 2) could be (1)(23)(45), (2)(35)(14), and so on. This cycle type of a permutation is exactly the level of detail that conjugacy measures in  $S_n$ : two permutations in  $S_n$  are conjugate precisely when they have the same cycle type. Let's understand how this works in an example first.

**Example 5.5.** We consider two permutations in  $S_5$  of cycle type (2, 3):

$$\pi_1 = (24)(153), \quad \pi_2 = (13)(425).$$

To conjugate  $\pi_1$  to  $\pi_2$ , let  $\sigma$  be the permutation in  $S_5$  that sends the terms appearing in  $\pi_1$  to the terms appearing in  $\pi_2$  in exactly the same order:  $\sigma = \begin{pmatrix} 24153 \\ 13425 \end{pmatrix} = (14352)$ . Then

$$\sigma\pi_1\sigma^{-1} = \sigma(24)(153)\sigma^{-1} = \sigma(24)\sigma^{-1}\sigma(153)\sigma^{-1} = (\sigma(2)\sigma(4))(\sigma(1)\sigma(5)\sigma(3)) = (13)(425),$$

so  $\sigma\pi_1\sigma^{-1} = \pi_2$ .

If we had written  $\pi_1$  and  $\pi_2$  differently, say as

$$\pi_1 = (42)(531), \quad \pi_2 = (13)(542),$$

then  $\pi_2 = \sigma\pi_1\sigma^{-1}$  where  $\sigma = \begin{pmatrix} 42531 \\ 13542 \end{pmatrix} = (1234)$ .

**Lemma 5.6.** *If  $\pi_1$  and  $\pi_2$  are disjoint permutations in  $S_n$ , then  $\sigma\pi_1\sigma^{-1}$  and  $\sigma\pi_2\sigma^{-1}$  are disjoint permutations for all  $\sigma \in S_n$ .*

*Proof.* Being disjoint means no number is moved by both  $\pi_1$  and  $\pi_2$ . That is, there is no  $i$  such that  $\pi_1(i) \neq i$  and  $\pi_2(i) \neq i$ . If  $\sigma\pi_1\sigma^{-1}$  and  $\sigma\pi_2\sigma^{-1}$  are not disjoint, then they both move some number, say  $j$ . Then (check!)  $\sigma^{-1}(j)$  is moved by both  $\pi_1$  and  $\pi_2$ , which is a contradiction.  $\square$

**Theorem 5.7.** *Two permutations in  $S_n$  are conjugate if and only if they have the same cycle type.*

*Proof.* Pick  $\pi \in S_n$ . Write  $\pi$  as a product of disjoint cycles. By Theorem 3.5 and Lemma 5.6,  $\sigma\pi\sigma^{-1}$  will be a product of the  $\sigma$ -conjugates of the disjoint cycles for  $\pi$ , and these  $\sigma$ -conjugates are *disjoint* cycles with the same respective lengths. Therefore  $\sigma\pi\sigma^{-1}$  has the same cycle type as  $\pi$ .

<sup>1</sup>A more descriptive label might be “disjoint cycle structure”, but the standard term is “cycle type”.



For the converse direction, we need to explain why permutations  $\pi_1$  and  $\pi_2$  with the same cycle type are conjugate. Suppose the cycle type is  $(m_1, m_2, \dots)$ . Then

$$\pi_1 = \underbrace{(a_1 \ a_2 \ \dots \ a_{m_1})}_{m_1 \text{ terms}} \underbrace{(a_{m_1+1} \ a_{m_1+2} \ \dots \ a_{m_1+m_2})}_{m_2 \text{ terms}} \cdots$$

and

$$\pi_2 = \underbrace{(b_1 \ b_2 \ \dots \ b_{m_1})}_{m_1 \text{ terms}} \underbrace{(b_{m_1+1} \ b_{m_1+2} \ \dots \ b_{m_1+m_2})}_{m_2 \text{ terms}} \cdots,$$

where the cycles here are disjoint. To carry  $\pi_1$  to  $\pi_2$  by conjugation in  $S_n$ , define the permutation  $\sigma \in S_n$  by:  $\sigma(a_i) = b_i$  for all  $i$ . Then  $\sigma\pi_1\sigma^{-1} = \pi_2$  by Theorems 3.5 and 5.4. (This is exactly the method used to find  $\sigma$  in Example 5.5.)  $\square$

Since the conjugacy class of a permutation in  $S_n$  is determined by its cycle type, which is a certain partition of  $n$ , the number of conjugacy classes in  $S_n$  is the number of partitions of  $n$ . The number of partitions of  $n$  is denoted  $p(n)$ . Here is a table of some values. Check the numbers at the start of the table for  $n \leq 6$  agree with the number of conjugacy classes in the tables at the start of this section.

$n$	1	2	3	4	5	6	7	8	9	10	11	12	13	14
$p(n)$	1	2	3	5	7	11	15	22	30	42	56	77	101	135

The function  $p(n)$  grows quickly, *e.g.*,  $p(100) = 190,569,292$ .

Using Theorem 5.7, there is a type of converse to Theorem 3.2: although elements of equal order in a group need not be conjugate in the group, conjugacy will occur by working in a larger group.

**Corollary 5.8.** *Each finite group  $G$  can be embedded in a larger group where all elements of  $G$  with equal order become conjugate.*

*Proof.* From Cayley's theorem we can embed  $G$  into a symmetric group by associating to each  $g \in G$  the permutation  $\ell_g: G \rightarrow G$  where  $\ell_g(x) = gx$ . By labeling the elements of  $G$  as  $\{g_1, \dots, g_n\}$  we can make each permutation of  $G$  look like a permutation of  $\{1, \dots, n\}$ , and that makes the mapping  $g \mapsto \ell_g$  an injective homomorphism  $G \rightarrow S_n$ , where  $n = |G|$ .

Let  $g \in G$  have order  $m$ , so  $m \mid n$ . Left multiplication of  $g$  on  $G$ , as a permutation of  $G$ , is a product of disjoint  $m$ -cycles  $(x \ gx \ g^2x \ \dots \ g^{m-1}x)$ . The cycle decomposition of  $\ell_g$  will have  $|G|/m$  disjoint  $m$ -cycles. Therefore the cycle type of the permutation  $\ell_g$  in  $S_n$  depends on  $g$  only through its order  $m$ . For another element  $g'$  in  $G$  with order  $m$ ,  $\ell_{g'}$  as a permutation of  $G$  has the same cycle type as  $\ell_g$ , so  $\ell_g$  and  $\ell_{g'}$  are conjugate in  $S_n$  by Theorem 5.7.  $\square$

**Example 5.9.** In the group  $G = \mathbf{Z}/(10)$ , 2 and 4 have order 5 and are not conjugate in  $G$  since  $G$  is abelian. When we view  $G$  in  $S_{10}$  using Cayley's theorem,  $\ell_2 = (02468)(13579)$  and  $\ell_4 = (04826)(15937)$ . These are conjugate in  $S_{10}$ :  $\ell_4 = \sigma\ell_2\sigma^{-1}$ , where  $\sigma = \begin{pmatrix} 0123456789 \\ 0145892367 \end{pmatrix} = (2486)(3597)$ .

Corollary 5.8 is also true for infinite  $G$ , by using infinite symmetric groups in the proof.

Let's now look at conjugacy classes in  $A_n$ . If  $\pi$  is an even permutation, then  $\sigma\pi\sigma^{-1}$  is also even, so a conjugacy class in  $S_n$  that contains one even permutation contains only even permutations. However, two permutations  $\pi_1$  and  $\pi_2$  in  $A_n$  can have the same cycle type (and thus be conjugate in the larger group  $S_n$ ) while being non-conjugate in  $A_n$ . The point is that we might be able to get  $\pi_2 = \sigma\pi_1\sigma^{-1}$  for some  $\sigma \in S_n$  without being able to do this for  $\sigma \in A_n$ .

**Example 5.10.** The 3-cycles (123) and (132) in  $A_3$  are conjugate in  $S_3$ :  $(23)(123)(23)^{-1} = (132)$ . However, (123) and (132) are not conjugate in  $A_3$  because  $A_3$  is abelian: an element of  $A_3$  is conjugate in  $A_3$  only to itself.

**Example 5.11.** The 3-cycle (123) and its inverse (132) are conjugate in  $S_4$  (by (23)) but they are not conjugate in  $A_4$ . To see this, let's determine all possible  $\sigma \in S_4$  that conjugate (123) to (132). For  $\sigma \in S_4$ , the condition  $\sigma(123)\sigma^{-1} = (132)$  is the same as  $(\sigma(1)\sigma(2)\sigma(3)) = (132)$ . There are three possibilities:

- $\sigma(1) = 1$ , so  $\sigma(2) = 3$  and  $\sigma(3) = 2$ , and necessarily  $\sigma(4) = 4$ . Thus  $\sigma = (23)$ .
- $\sigma(1) = 3$ , so  $\sigma(2) = 2$  and  $\sigma(3) = 1$ , and necessarily  $\sigma(4) = 4$ . Thus  $\sigma = (13)$ .
- $\sigma(1) = 2$ , so  $\sigma(2) = 1$  and  $\sigma(3) = 3$ , and necessarily  $\sigma(4) = 4$ . Thus  $\sigma = (12)$ .

Therefore the only possible  $\sigma$ 's are transpositions, which are not in  $A_4$ .

While it would be nice if conjugacy classes in  $A_n$  are determined by cycle type as in  $S_n$ , we have seen that this is false: (123) and (132) are not conjugate in  $A_3$  or  $A_4$ . How does a conjugacy class of even permutations in  $S_n$  break up when thinking about conjugacy classes in  $A_n$ ? There are two possibilities: the conjugacy class stays as a single conjugacy class within  $A_n$  or it breaks up into two conjugacy classes of equal size in  $A_n$ . A glance at the earlier tables of conjugacy classes in  $A_n$  with small  $n$  shows this happening. For instance,

- there is one class of 8 3-cycles in  $S_4$ , but two classes of 4 3-cycles in  $A_4$ ,
- there is one class of 24 5-cycles in  $S_5$ , but two classes of 12 5-cycles in  $A_5$ ,
- there is one class of 144 5-cycles in  $S_6$ , but two classes of 72 5-cycles in  $A_6$ .

A rule that describes when each possibility occurs is as follows, but a proof is omitted.

**Theorem 5.12.** *For  $\pi \in A_n$ , its conjugacy class in  $S_n$  remains as a single conjugacy class in  $A_n$  or it breaks into two conjugacy classes in  $A_n$  of equal size. The conjugacy class breaks up if and only if the lengths in the cycle type of  $\pi$  are distinct odd numbers.*

Here is a table showing the cycle types in  $A_n$  that fall into two conjugacy classes for  $4 \leq n \leq 14$ . For example, the permutations in  $A_6$  of cycle type (1, 5) but *not* (3, 3) fall into two conjugacy classes and the permutations in  $A_8$  of cycle type (1, 7) and (3, 5) each fall into two conjugacy classes.

$n$	4	5	6	7	8	9	10	11	12	13	14
Cycle type in $A_n$	(1,3)	(5)	(1,5)	(7)	(1,7)	(9)	(1,9)	(11)	(1,11)	(13)	(1,13)
					(3,5)	(1,3,5)	(3,7)	(1,3,7)	(3,9)	(1,3,9)	(3,11)
									(5,7)	(1,5,7)	(5,9)

The following table lists the number  $c(n)$  (nonstandard notation) of conjugacy classes in  $A_n$  for small  $n$ .

$n$	1	2	3	4	5	6	7	8	9	10	11	12	13	14
$c(n)$	1	1	3	4	5	7	9	14	18	24	31	43	55	72

## 6. CENTRALIZERS AND THE CLASS EQUATION

We saw in Theorem 3.4 that different conjugacy classes do not overlap. Thus, they provide a way of covering the group by disjoint sets. This is analogous to the left cosets of a subgroup providing a disjoint covering of the group.

For  $g \in G$ , let  $K_g$  denote its conjugacy class in  $G$ :

$$K_g = \{xgx^{-1} : x \in G\}.$$

If the different conjugacy classes are  $K_{g_1}, K_{g_2}, \dots, K_{g_r}$ , then

$$(6.1) \quad |G| = |K_{g_1}| + |K_{g_2}| + \dots + |K_{g_r}|.$$

Equation (6.1) plays the role for conjugacy classes in  $G$  that the formula  $|G| = |H|[G : H]$  plays for cosets of  $H$  in  $G$ .

Let's see how (6.1) looks for some groups from Section 2.

**Example 6.1.** For  $G = S_3$ , (6.1) says

$$6 = 1 + 2 + 3.$$

**Example 6.2.** For  $G = D_4$ ,

$$8 = 1 + 1 + 2 + 2 + 2.$$

**Example 6.3.** For  $G = Q_8$ ,

$$8 = 1 + 1 + 2 + 2 + 2.$$

**Example 6.4.** For  $G = A_4$ ,

$$12 = 1 + 3 + 4 + 4.$$

The reason (6.1) is important is that  $|K_{g_i}|$  divides the size of the group. We saw this earlier in examples. Now we will prove it in general.

**Theorem 6.5.** *If  $G$  is a finite group then each conjugacy class in  $G$  has size dividing  $|G|$ .*

Theorem 6.5 is not an immediate consequence of Lagrange's theorem, because conjugacy classes are *not* subgroups. For example, no conjugacy class contains the identity except for the one-element conjugacy class containing the identity by itself. However, while a conjugacy class is not a subgroup, its size does equal the *index* of a subgroup, and that will explain why its size divides the size of the group.

**Definition 6.6.** For a group  $G$ , its *center*  $Z(G)$  is the set of elements of  $G$  commuting with everything:

$$Z(G) = \{g \in G : gx = xg \text{ for all } x \in G\}.$$

For  $g \in G$ , its *centralizer*  $Z(g)$  is the set of elements of  $G$  commuting with  $g$ :

$$Z(g) = \{x \in G : xg = gx\}.$$

The notation  $Z$  comes from German: center is Zentrum and centralizer is Zentralisator. Some English language books use the letter  $C$ , so  $C(G) = Z(G)$  and  $C(g) = Z(g)$ . The center of the group and the centralizer of each element of the group are subgroups. The connection between them is the center is the intersection of all the centralizers:  $Z(G) = \bigcap_{g \in G} Z(g)$ .

**Theorem 6.7.** *For each  $g \in G$ , its conjugacy class has the same size as the index of its centralizer:*

$$|\{xgx^{-1} : x \in G\}| = [G : Z(g)].$$

*Proof.* Consider the function  $f: G \rightarrow K_g$  where  $f(x) = xgx^{-1}$ . This function is onto, since by definition every element of  $K_g$  is  $xgx^{-1}$  for some  $x \in G$ . We will now examine when  $f$  takes the same value at elements of  $G$ .

For  $x$  and  $x'$  in  $G$ , we have  $xgx^{-1} = x'gx'^{-1}$  if and only if

$$gx^{-1}x' = x^{-1}x'g.$$

Therefore  $x^{-1}x'$  commutes with  $g$ , *i.e.*,  $x^{-1}x' \in Z(g)$ , so  $x' \in xZ(g)$ . Although  $x$  and  $x'$  may be different, they lie in the same left coset of  $Z(g)$ :

$$(6.2) \quad f(x) = f(x') \implies xZ(g) = x'Z(g).$$

Conversely, suppose  $xZ(g) = x'Z(g)$ . Then  $x = x'z$  for some  $z \in Z(g)$ , so  $zg = gz$ . Therefore  $x$  and  $x'$  conjugate  $g$  in the same way:

$$\begin{aligned} f(x) &= xgx^{-1} \\ &= (x'z)g(x'z)^{-1} \\ &= x'zgz^{-1}x'^{-1} \\ &= x'gzx'^{-1} \\ &= x'gx'^{-1} \\ &= f(x'). \end{aligned}$$

Since we have shown that the converse of (6.2) is true, the function  $f: G \rightarrow K_g$  takes the same value at two elements precisely when they are in the same left coset of  $Z(g)$ . Therefore the number of different values of  $f$  is the number of different left cosets of  $Z(g)$  in  $G$ , and by definition that is the index  $[G : Z(g)]$ . Since  $f$  is surjective, we conclude that  $|K_g| = [G : Z(g)]$ .  $\square$

Now we can prove Theorem 6.5.

*Proof.* By Theorem 6.7, the size of the conjugacy class of  $g$  is the index  $[G : Z(g)]$ , which divides  $|G|$ .  $\square$

Returning to (6.1), we rewrite it in the form

$$(6.3) \quad |G| = \sum_{i=1}^r [G : Z(g_i)] = \sum_{i=1}^r \frac{|G|}{|Z(g_i)|}.$$

The conjugacy classes of size 1 are exactly those containing elements of the center of  $G$  (*i.e.*, those  $g_i$  such that  $Z(g_i) = G$ ). Combining all of these 1's into a single term, we get

$$(6.4) \quad |G| = |Z(G)| + \sum_{i'} \frac{|G|}{|Z(g_{i'})|},$$

where the sum is now carried out only over those conjugacy classes  $K_{g_{i'}}$  with more than one element. In the terms of this sum,  $|Z(g_{i'})| < |G|$ . Equation (6.4) is called the *class equation*. The difference between the class equation and (6.1) is that we have combined the terms contributing to the center of  $G$  into a single term.

Here is a good application of the class equation.

**Theorem 6.8.** *When  $G$  is a nontrivial finite  $p$ -group it has a nontrivial center: some element of  $G$  other than the identity commutes with every element of  $G$ .*

*Proof.* Let  $|G| = p^n$ , where  $n > 0$ . Consider a term  $[G : Z(g_{i'})]$  in the class equation, where  $g_{i'}$  does not lie in  $Z(G)$ . Then  $Z(g_{i'}) \neq G$ , so the index  $[G : Z(g_{i'})]$  is a factor of  $|G|$  other than 1. It is one of  $\{p, p^2, \dots, p^n\}$ , and hence is *divisible by  $p$* . In the class equation, all terms in the sum over  $i'$  are multiples of  $p$ .

Also, the left side of the class equation is a multiple of  $p$ , since  $|G| = p^n$ . So the class equation forces  $p \mid |Z(G)|$ . Since the center contains the identity, and has size divisible by  $p$ , it must contain non-identity elements as well.  $\square$

With a little extra work we can generalize Theorem 6.8.

**Theorem 6.9.** *If  $G$  is a nontrivial finite  $p$ -group and  $N$  is a nontrivial normal subgroup of  $G$  then  $N \cap Z(G) \neq \{e\}$ .*

*Proof.* Since  $N$  is a normal subgroup of  $G$ , a conjugacy class in  $G$  that meets  $N$  lies entirely inside of  $N$  (that is, if  $g \in N$  then  $xgx^{-1} \in N$  for all  $x \in G$ ). Let  $K_{g_1}, \dots, K_{g_s}$  be the different conjugacy classes of  $G$  that lie inside  $N$ , so

$$(6.5) \quad |N| = |K_{g_1}| + \dots + |K_{g_s}|.$$

(Note that elements of  $N$  can be conjugate in  $G$  without being conjugate in  $N$ , so breaking up  $N$  into its  $G$ -conjugacy classes in (6.5) is a coarser partitioning of  $N$  than breaking it into  $N$ -conjugacy classes.) The left side of (6.5) is a power of  $p$  greater than 1. Each term on the right side is a conjugacy class in  $G$ , so  $|K_{g_i}| = [G : Z(g_i)]$ , where  $Z(g_i)$  is the centralizer of  $g_i$  in  $G$ . This index is a power of  $p$  greater than 1 except when  $g_i \in Z(G)$ , in which case  $|K_{g_i}| = 1$ . The  $g_i$ 's in  $N$  with  $|K_{g_i}| = 1$  are elements of  $N \cap Z(G)$ . Therefore if we reduce (6.5) modulo  $p$  we get

$$0 \equiv |N \cap Z(G)| \pmod{p},$$

so  $|N \cap Z(G)|$  is divisible by  $p$ . Since  $|N \cap Z(G)| \geq 1$  the intersection  $N \cap Z(G)$  contains a non-identity term.  $\square$

**Remark 6.10.** The finiteness assumption in Theorem 6.8 is important. There are infinite  $p$ -groups with trivial center! Here is an example. Consider the set  $G$  of infinite mod  $p$  square matrices  $\begin{pmatrix} M & O \\ O & I_\infty \end{pmatrix}$  where  $I_\infty$  is an infinite identity matrix and  $M$  is a finite upper triangular square matrix of the form

$$\begin{pmatrix} 1 & a_{12} & \cdots & a_{1n} \\ 0 & 1 & \cdots & a_{2n} \\ 0 & 0 & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}$$

where there are 1's on the main diagonal and the entries  $a_{ij}$  above the main diagonal are in  $\mathbf{Z}/(p)$ . Because each row or column of a matrix in  $G$  has only finitely many nonzero elements, matrix multiplication in  $G$  makes sense. To show  $G$  is a group under matrix multiplication, by borrowing the upper left 1 in  $I_\infty$  we can write

$$\begin{pmatrix} M & O \\ O & I_\infty \end{pmatrix} = \begin{pmatrix} M & O & O \\ O & 1 & O \\ O & O & I_\infty \end{pmatrix}$$

and thereby view the infinite matrix as having an  $(n+1) \times (n+1)$  upper left part instead of an  $n \times n$  upper left part. In this way all pairs of matrices  $\begin{pmatrix} M & O \\ O & I_\infty \end{pmatrix}$  and  $\begin{pmatrix} N & O \\ O & I_\infty \end{pmatrix}$  in  $G$  can be considered to have  $M$  and  $N$  of the same size. Then we obtain a block multiplication formula (check!)  $\begin{pmatrix} M & O \\ O & I_\infty \end{pmatrix} \begin{pmatrix} N & O \\ O & I_\infty \end{pmatrix} = \begin{pmatrix} MN & O \\ O & I_\infty \end{pmatrix}$ . Since the  $n \times n$  upper triangular mod  $p$  matrices with 1's on the main diagonal form a group, it now follows that  $G$  is a group. Since  $M$  has  $p$ -power order, every element of  $G$  has  $p$ -power order. Thus  $G$  is an "infinite  $p$ -group."

To show the center of  $G$  is trivial, a non-identity element of  $G$  has the form  $\begin{pmatrix} M & O \\ O & I_\infty \end{pmatrix}$ , where  $M$  is  $n \times n$  for some  $n$  and  $M \neq I_n$ . We have the following equations in  $2n \times 2n$  matrices:

$$\begin{pmatrix} M & O \\ O & I_n \end{pmatrix} \begin{pmatrix} I_n & I_n \\ O & I_n \end{pmatrix} = \begin{pmatrix} M & M \\ O & I_n \end{pmatrix},$$

$$\begin{pmatrix} I_n & I_n \\ O & I_n \end{pmatrix} \begin{pmatrix} M & O \\ O & I_n \end{pmatrix} = \begin{pmatrix} M & I_n \\ O & I_n \end{pmatrix}.$$

These are not equal since  $M \neq I_n$ . Now embed the  $2n \times 2n$  matrices  $A = \begin{pmatrix} M & O \\ O & I_n \end{pmatrix}$  and  $B = \begin{pmatrix} I_n & I_n \\ O & I_n \end{pmatrix}$  in  $G$  as  $\begin{pmatrix} A & O \\ O & I_\infty \end{pmatrix}$  and  $\begin{pmatrix} B & O \\ O & I_\infty \end{pmatrix}$ . These do not commute. Note  $\begin{pmatrix} A & O \\ O & I_\infty \end{pmatrix} = \begin{pmatrix} M & O \\ O & I_\infty \end{pmatrix}$  in  $G$ , so  $\begin{pmatrix} M & O \\ O & I_\infty \end{pmatrix} \notin Z(G)$ .

The following corollary is the standard first application of Theorem 6.8.

**Corollary 6.11.** *For all primes  $p$ , every group of order  $p^2$  is abelian. More precisely, a group of order  $p^2$  is isomorphic to  $\mathbf{Z}/(p^2)$  or to  $\mathbf{Z}/(p) \times \mathbf{Z}/(p)$ .*

*Proof.* Let  $G$  be such a group. By Lagrange, the order of a non-identity element is  $p$  or  $p^2$ .

If there is an element of  $G$  with order  $p^2$ , then  $G$  is cyclic and therefore isomorphic to  $\mathbf{Z}/(p^2)$  (in many ways). We may henceforth assume  $G$  has no element of order  $p^2$ . That means every non-identity element of  $G$  has order  $p$ .

From Theorem 6.8, there is a non-identity element in the center of  $G$ . Call it  $a$ . Since  $a$  has order  $p$ ,  $\langle a \rangle$  is not all of  $G$ . Choose  $b \in G - \langle a \rangle$ . Then  $b$  also has order  $p$ . We are going to show powers of  $a$  and powers of  $b$  provide an isomorphism of  $G$  with  $\mathbf{Z}/(p) \times \mathbf{Z}/(p)$ . Let  $f: \mathbf{Z}/(p) \times \mathbf{Z}/(p) \rightarrow G$  by

$$f(i, j) = a^i b^j.$$

This is well-defined since  $a$  and  $b$  have order  $p$ . It is a homomorphism since powers of  $a$  are in the center:

$$\begin{aligned} f(i, j)f(i', j') &= (a^i b^j)(a^{i'} b^{j'}) \\ &= a^i a^{i'} b^j b^{j'} \\ &= a^{i+i'} b^{j+j'} \\ &= f(i+i', j+j') \\ &= f((i, j) + (i', j')). \end{aligned}$$

The kernel is trivial: if  $f(i, j) = e$  then  $a^i = b^{-j}$ . This is a common element of  $\langle a \rangle \cap \langle b \rangle$ , which is trivial. Therefore  $a^i = b^j = e$ , so  $i = j = 0$  in  $\mathbf{Z}/(p)$ .

Since  $f$  has trivial kernel it is injective. The domain and target have the same size, so  $f$  is surjective and thus is an isomorphism.  $\square$

**Corollary 6.12.** *A finite  $p$ -group  $\neq \{e\}$  has a normal subgroup of order  $p$ .*

*Proof.* Let  $G$  be a finite  $p$ -group with  $|G| > 1$ . By Theorem 6.8,  $Z(G)$  is a nontrivial  $p$ -group. Pick  $g \in Z(G)$  with  $g \neq e$ . The order of  $g$  is  $p^r$  for some  $r \geq 1$ . Therefore  $g^{p^{r-1}}$  has order  $p$ , so  $Z(G)$  contains a subgroup of order  $p$ , which must be normal in  $G$  since every subgroup of  $Z(G)$  is a normal subgroup of  $G$ .  $\square$

We can bootstrap Corollary 6.12 to non-prime sizes by inducting on a stronger hypothesis.

**Corollary 6.13.** *If  $G$  is a nontrivial finite  $p$ -group with size  $p^n$  then there is a normal subgroup of size  $p^j$  for every  $j = 0, 1, \dots, n$ .*

*Proof.* We argue by induction on  $n$ . The result is clear if  $n = 1$ . Suppose  $n \geq 2$  and the theorem is true for  $p$ -groups of size  $p^{n-1}$ . If  $|G| = p^n$  then it has a normal subgroup  $N$  of size  $p$  by the preceding corollary. Then  $|G/N| = p^{n-1}$ , so for  $0 \leq j \leq n-1$  there is a normal subgroup of  $G/N$  with size  $p^j$ . The pullback of this subgroup to  $G$  is normal and has size  $p^j \cdot |N| = p^{j+1}$ .  $\square$

**Example 6.14.** Let  $G = D_4$ . Its subgroups of size 2 are  $\langle s \rangle$ ,  $\langle rs \rangle$ ,  $\langle r^2s \rangle$ ,  $\langle r^3s \rangle$ , and  $\langle r^2 \rangle$ . The last one is normal. The subgroups of size 4 are  $\langle r \rangle$  and  $\langle r^2, s \rangle$ . Both are normal.

APPENDIX A. CONJUGACY IN PLANE GEOMETRY

We will show that all reflections in  $\mathbf{R}^2$  are conjugate to reflection across the  $x$ -axis in an appropriate group of transformations of the plane.

**Definition A.1.** An *isometry* of  $\mathbf{R}^2$  is a function  $f: \mathbf{R}^2 \rightarrow \mathbf{R}^2$  that preserves distances: for all points  $P$  and  $Q$  in  $\mathbf{R}^2$ , the distance between  $f(P)$  and  $f(Q)$  is the same as the distance between  $P$  and  $Q$ .

Isometries of  $\mathbf{R}^2$  include: translations, rotations, and reflections.<sup>2</sup> Isometries are invertible (this requires proof, or include it in the definition if you want to be lazy about it), and under composition isometries form a group.

There are two ways to describe points of the plane algebraically, using vectors or complex numbers. We will work with points as complex numbers. The point  $(a, b)$  is considered as the complex number  $a + bi$ . We measure the distance to  $a + bi$  from 0 with the absolute value

$$|a + bi| = \sqrt{a^2 + b^2},$$

and the distance between  $a + bi$  and  $c + di$  is the absolute value of their difference:

$$|(a + bi) - (c + di)| = \sqrt{(a - c)^2 + (b - d)^2}.$$

To each complex number  $z = a + bi$ , we have its complex conjugate  $\bar{z} = a - bi$ . By an explicit calculation, complex conjugation respects sums and products:

$$\overline{z + z'} = \bar{z} + \bar{z'}, \quad \overline{zz'} = \bar{z}\bar{z'}.$$

Two important algebraic properties of the absolute value on  $\mathbf{C}$  are its behavior on products and on complex conjugates:

$$|zz'| = |z||z'|, \quad |\bar{z}| = |z|.$$

In particular, if  $|w| = 1$  then  $|wz| = |z|$ .

An example of a reflection across a line in the plane is complex conjugation:

$$s(z) = \bar{z}.$$

This is reflection across the  $x$ -axis. It preserves distance:

$$|s(z) - s(z')| = |\bar{z} - \bar{z}'| = \overline{|z - z'|} = |z - z'|.$$

We will compare this reflection with the reflection across another line, first treating other lines through the origin and then treating lines that may not pass through the origin.

Pick a line through the origin that makes an angle, say  $\theta$ , with respect to the positive  $x$ -axis. We can rotate the  $x$ -axis onto that line by rotating the  $x$ -axis counterclockwise around the origin through an angle of  $\theta$ . A rotation around the origin, in terms of complex numbers, is multiplication by the number  $\cos \theta + i \sin \theta$ , which has absolute value 1. Let's denote counterclockwise rotation around the origin by  $\theta$  by  $r_\theta$ :

$$(A.1) \quad r_\theta(z) = (\cos \theta + i \sin \theta)z, \quad |\cos \theta + i \sin \theta| = 1.$$

---

<sup>2</sup>A full description of isometries of  $\mathbf{R}^2$  includes glide reflections. See <https://kconrad.math.uconn.edu/blurbs/grouptheory/isometrycp.pdf> or <https://kconrad.math.uconn.edu/blurbs/grouptheory/isometryR2.pdf>.

Every rotation  $r_\theta$  preserves distances:

$$|r_\theta(z) - r_\theta(z')| = |(\cos \theta + i \sin \theta)(z - z')| = |(\cos \theta + i \sin \theta)||z - z'| = |z - z'|.$$

Composing rotations around the origin amounts to adding angles:  $r_\theta \circ r_\varphi = r_{\theta+\varphi}$ . In particular,  $r_\theta^{-1} = r_{-\theta}$  since  $r_\theta \circ r_{-\theta} = r_0$ , which is the identity ( $r_0(z) = z$ ).

Now let's think about some reflections besides complex conjugation. Let  $s_\theta$  be the reflection across the line through the origin making an angle of  $\theta$  with the positive  $x$ -axis. (In particular, complex conjugation is  $s_0$ .) Draw some pictures to convince yourself visually the reflection  $s_\theta$  is the composite of

- rotation of the plane by an angle of  $-\theta$  to bring the line of reflection onto the  $x$ -axis,
- reflection across the  $x$ -axis,
- rotation of the plane by  $\theta$  to return the line to its original position.

This says

$$(A.2) \quad s_\theta = r_\theta s r_{-\theta} = r_\theta s r_\theta^{-1}.$$

So we see, in this algebraic formula, that a reflection across each line through the origin is *conjugate*, in the group of isometries of the plane, to reflection across the  $x$ -axis. The conjugating isometry is the rotation  $r_\theta$  that takes the line through the origin at angle  $\theta$  to the  $x$ -axis.

In order to compare complex conjugation to reflection across an arbitrary line, which need not pass through the origin, we bring in additional isometries: translations. A translation in the plane can be viewed as adding a particular complex number, say  $w$ , to every complex number:  $t_w(z) = z + w$ . This is an isometry since

$$|t_w(z) - t_w(z')| = |(z + w) - (z' + w)| = |z - z'|.$$

Note  $t_w \circ t_{w'} = t_{w+w'}$ , and the inverse of  $t_w$  is  $t_{-w}$ :  $t_w^{-1} = t_{-w}$ .

In order to describe reflection across an arbitrary line in terms of complex conjugation, we need to describe an arbitrary line. A line makes a definite angle with respect to the positive  $x$ -direction (how far it tilts). Call that angle  $\theta$ . Now pick a point on the line. Call it, say,  $w$ . Our line is the only line in the plane that passes through  $w$  at an angle of  $\theta$  relative to the positive  $x$ -direction.

We can carry out reflection across this line in terms of reflection across the line parallel line to it through the origin by using translations, in 3 steps:

- translate *back* by  $w$  (that is, apply  $t_{-w}$ ) to carry the original line to a line through the origin at the same angle  $\theta$ ,
- reflect across this line through the origin (apply  $s_\theta$ ),
- translate by  $w$  to return the line to its original position (apply  $t_w$ ).

Putting this all together, with (A.2), reflection across the line through  $w$  that makes an angle of  $\theta$  with the positive  $x$ -direction is the composite

$$(A.3) \quad t_w s_\theta t_{-w} = t_w (r_\theta s r_\theta^{-1}) t_w^{-1} = t_w r_\theta s (t_w r_\theta)^{-1}.$$

This is a *conjugate* of complex conjugation  $s$  in the group of isometries in the plane.

Let's summarize what we have shown.

**Theorem A.2.** *In the group of isometries of the plane, reflection across a line is conjugate to reflection across the  $x$ -axis.*



**Example A.3.** Reflection across the horizontal line  $y = b$  corresponds to  $\theta = 0$  and  $w = bi$ . That is, this reflection is  $t_{bi}st_{-bi}$ : translate down by  $b$ , reflect across the  $x$ -axis, and then translate up by  $b$ .

APPENDIX B. BOUNDING A GROUP'S SIZE BY NUMBER OF CONJUGACY CLASSES

Up to isomorphism there obviously are only finitely many finite groups with a given size. What might be more surprising is that, up to isomorphism, there are only finitely many finite groups with a given number of conjugacy classes. The following theorem was proved by Landau [8] in 1903 as an application of the class equation.

**Theorem B.1.** *The size of a finite group can be bounded above from knowing the number of its conjugacy classes.*

*Proof.* When there is only one conjugacy class, the group is trivial. Now fix a positive integer  $k > 1$  and let  $G$  be a finite group with  $k$  conjugacy classes represented by  $g_1, \dots, g_k$  (this includes  $g_i$ 's in the center). We exploit the class equation, written as

$$(B.1) \quad |G| = \sum_{i=1}^k \frac{|G|}{|Z(g_i)|}.$$

Dividing (B.1) by  $|G|$ ,

$$(B.2) \quad 1 = \frac{1}{n_1} + \frac{1}{n_2} + \dots + \frac{1}{n_k},$$

where  $n_i = |Z(g_i)|$ . Note each  $n_i$  exceeds 1 when  $G$  is nontrivial. We write the  $n_i$ 's in increasing order, with possible repetitions:

$$(B.3) \quad n_1 \leq n_2 \leq \dots \leq n_k.$$

Since each  $n_i$  is at least as large as  $n_1$ , (B.2) implies

$$1 \leq \frac{k}{n_1},$$

so

$$(B.4) \quad n_1 \leq k.$$

Then, using  $n_i \geq n_2$  for  $i \geq 2$ ,

$$1 \leq \frac{1}{n_1} + \frac{k-1}{n_2}.$$

Thus  $1 - 1/n_1 \leq (k-1)/n_2$ , so

$$(B.5) \quad n_2 \leq \frac{k-1}{1 - 1/n_1}.$$

By induction,

$$(B.6) \quad n_m \leq \frac{k+1-m}{1 - (\frac{1}{n_1} + \dots + \frac{1}{n_{m-1}})}$$

for  $m \geq 2$ .

Since (B.4) bounds  $n_1$  from above by  $k$  and (B.6) bounds each of  $n_2, \dots, n_k$  from above in terms of earlier  $n_i$ 's, there are only a finite number of  $k$ -tuples  $(n_1, \dots, n_k)$ . The  $k$ -tuples among these that satisfy (B.2) can be tabulated. The largest value of  $n_k$  is  $|G|$  (since 1 has

centralizer  $G$ ), so the solution with the largest value for  $n_k$  gives an upper bound on the size of a finite group with  $k$  conjugacy classes.  $\square$

The number of conjugacy classes in a finite group is called the *class number* of that group.<sup>3</sup> Since the identity element of a group is its own conjugacy class, the only finite group with class number 1 is the trivial group. Burnside [4, p. 461–462] and Miller [9] in 1911 independently determined all the finite groups (up to isomorphism) with class number 2, 3, 4, and 5. Poland [10] determined the groups with class number 6 and 7. Currently a complete list of groups with class number  $k$  is known for  $k \leq 14$  [12]. The results through  $k = 7$  are given below.

A group of order  $n$  has at most  $n$  conjugacy classes, and there are  $n$  conjugacy classes if and only if the group is abelian. Therefore a list of finite groups with  $k$  conjugacy classes will include all the abelian groups of order  $k$ . When  $n = 2m + 1$  is odd and at least 3, the dihedral group  $D_n$  has  $m + 2$  conjugacy classes (they are  $\{1\}$ ,  $\{r^{\pm i}\}$  for  $1 \leq i \leq (n - 1)/2$ , and the reflections, so the number of conjugacy classes is  $1 + (n - 1)/2 + 1 = m + 3$ .) When  $k = m + 3$ ,  $2m + 1 = 2k - 5$ , so each integer  $k \geq 3$  is the class number of at least one nonabelian group, namely  $D_{2k-3}$ .

**Example B.2.** Taking  $k = 2$ , the only solution to (B.2) satisfying (B.4) and (B.6) is  $n_1 = 2$ ,  $n_2 = 2$ . Thus  $G \cong \mathbf{Z}/(2)$ . That is, the finite groups with exactly two conjugacy classes are cyclic of order 2. Note there are infinite groups with exactly two conjugacy classes, but they are not easy to describe: see <https://mathoverflow.net/questions/146799>.

**Example B.3.** When  $k = 3$ , the  $(n_1, n_2, n_3)$  satisfying (B.2), (B.3), (B.4), and (B.6) are  $(2, 4, 4)$ ,  $(2, 3, 6)$ , and  $(3, 3, 3)$ . Thus  $|G| \leq 6$  and  $S_3 \cong D_3$  has size 6 with 3 conjugacy classes while  $\mathbf{Z}/(3)$  has size 3 with 3 conjugacy classes.

**Example B.4.** When  $k = 4$ , there are 14 solutions, such as  $4, 4, 4, 4$  and  $2, 3, 7, 42$ . The second 4-tuple is actually the one with the largest value of  $n_4$ , so a finite group with 4 conjugacy classes has size at most 42. In actuality, the finite groups with 4 conjugacy classes are  $\mathbf{Z}/(4)$ ,  $(\mathbf{Z}/(2))^2$ ,  $D_{10}$ , and  $A_4$ , so the largest such finite group has order 12.

**Example B.5.** When  $k = 5$ , there are 148 solutions, and the largest  $n_5$  that occurs is 1806. The finite groups with 5 conjugacy classes are  $\mathbf{Z}/(5)$ ,  $Q_8$ ,  $D_4$ ,  $D_7$ ,  $\text{Aff}(\mathbf{Z}/(5))$ ,  $S_4$ ,  $A_5$ , and the nonabelian group of size 21, so the largest such size is  $|A_5| = 60$ .

**Example B.6.** The finite groups with 6 conjugacy classes are  $\mathbf{Z}/(6)$ ,  $D_6$ ,  $D_9$ ,  $\mathbf{Z}/(3) \rtimes \mathbf{Z}/(4)$ ,  $\mathbf{Z}/(3) \rtimes S_3$ ,  $\text{Aff}(\mathbf{F}_9)$ ,  $\{ \begin{pmatrix} a & b \\ 0 & 1 \end{pmatrix} : a, b \in \mathbf{F}_9, a^4 = 1 \}$ , and  $\text{PSL}_2(\mathbf{Z}/(7))$ . The largest size is  $|\text{PSL}_2(\mathbf{Z}/(7))| = 168$ .

In the semidirect products  $\mathbf{Z}/(3) \rtimes \mathbf{Z}/(4)$  and  $\mathbf{Z}/(3) \rtimes S_3$ , the action of  $\mathbf{Z}/(4)$  on  $\mathbf{Z}/(3)$  and  $S_3$  on  $\mathbf{Z}/(3)$  is by inversion through natural homomorphisms to  $\{\pm 1\}$ :  $k \bmod 4 \mapsto (-1)^k$  and  $\sigma \mapsto \text{sign}(\sigma)$ .

**Example B.7.** The 12 finite groups with 7 conjugacy classes are  $\mathbf{Z}/(7)$ ,  $D_8$ ,  $D_{11}$ ,  $S_5$ ,  $A_6$ ,  $\text{Aff}(\mathbf{Z}/(7))$ ,  $\text{SL}_2(\mathbf{Z}/(3))$ ,  $\{ \begin{pmatrix} a & b \\ 0 & 1 \end{pmatrix} : a, b \in \mathbf{Z}/(13), a^3 = 1 \}$ ,  $\{ \begin{pmatrix} a & b \\ 0 & 1 \end{pmatrix} : a, b \in \mathbf{Z}/(13), a^4 = 1 \}$ ,  $\{ \begin{pmatrix} a & b \\ 0 & 1 \end{pmatrix} : a, b \in \mathbf{Z}/(11), a^5 = 1 \}$ ,  $Q_{16}$  (generalized quaternion group of order 16),<sup>4</sup> and the subgroup of order 16 in  $\text{GL}_2(\mathbf{Z}/(3))$ . The group of largest order is  $|A_6| = 360$ .

<sup>3</sup>The term “class number” is used in algebraic number theory with an entirely different meaning: the size of the ideal class group of a number field.

<sup>4</sup>See <https://kconrad.math.uconn.edu/blurbs/grouptheory/genquat.pdf>.

For  $k \in \mathbf{Z}^+$ , all sufficiently large finite groups have more than  $k$  conjugacy classes by Theorem B.1. Let  $c(n)$ , for  $n \geq 1$ , be the fewest number of conjugacy classes among all groups of order  $n$ . so  $c(n) \rightarrow \infty$  as  $n \rightarrow \infty$ . Can this growth be quantified? In 1963, Brauer showed  $c(n) \geq \log_2 \log_2 n$  for  $n \geq 2$  (note  $\log_2 \log_2 2 = 0$ ). In 1992, Pyber [11] gave an exponential improvement by essentially removing one logarithm:  $c(n) \geq A \log_2(n)/(\log_2 \log_2 n)^8$  for some constant  $A > 0$ . His proof relies on the classification of finite simple groups. Almost 20 years later, Keller [7] in 2011 reduced the exponent in Pyber's lower bound from 8 to 7. In 2015, Baumeister, Maróti, and Tong-Viet [1] reduced the exponent to an arbitrary number greater than 3: for all  $\varepsilon > 0$  there is  $A_\varepsilon > 0$  such that  $c(n) \geq A_\varepsilon \log_2 n / (\log_2(\log_2 n))^{3+\varepsilon}$  for  $n \geq 3$ . Their proof, like the earlier work, uses the classification of finite simple groups. It is natural to guess that  $c(n) \geq A \log_2 n$  for some  $A > 0$ . Bertram [2] conjectured  $c(n) > \log_3 n$  for  $n \geq 1$ .

There is a logarithmic upper bound on  $c(n)$  for infinitely many  $n$ : for prime  $p$ , the set of mod  $p$  polynomials modulo  $x^{p+2}$  with constant term 0, which is

$$\{x + a_2x^2 + \cdots + a_{p+1}x^{p+1} \bmod x^{p+2} : a_i \in \mathbf{Z}/(p)\},$$

forms a finite group under composition (not under multiplication!)<sup>5</sup> with order  $p^p$  and turns out to have at most  $p^3$  conjugacy classes [5, Sect. 3.3]. For large  $p$ ,  $p^3$  is logarithmically smaller than  $p^p$ : setting  $n = p^p$ , so  $\log n = p \log p > p$  (for  $p \geq 3$ ), we have  $p^3 < (\log n)^3$ . Therefore  $c(n) < (\log n)^3$  when  $n = p^p$  for all odd primes  $p$ . This is not true for  $p = 2$ , since all groups of order 4 are abelian, so  $c(4) = 4 > 1 > (\log 2)^3$ .

For all  $n$ , a group of order  $n$  has at most  $n$  conjugacy classes, so  $c(n) \leq n$ , and  $c(n) = n$  if all groups of order  $n$  are abelian. That occurs infinitely often, e.g., when  $n$  is prime.

## REFERENCES

- [1] B. Baumeister, A. Maróti, and H. P. Tong-Viet, Finite groups have more conjugacy classes, *Forum Math.* **29** (2017), 259–275.
- [2] E. A. Bertram, New reductions and logarithmic lower bounds for the number of conjugacy classes infinite groups, *Bull. Aust. Math. Soc.* **87** (2013), 406–424.
- [3] A. Bensaïd and R. W. van der Waall, On finite groups all of whose elements of equal order are conjugate, *Simon Stevin* **65** (1991), 361–374.
- [4] W. Burnside, “Theory of groups of finite order,” 2nd ed., Cambridge Univ. Press, Cambridge, 1911. URL <https://catalog.hathitrust.org/Record/000419368>.
- [5] P. Etingof, On some properties of quantum doubles of finite groups, *J. Algebra* **394** (2013), 1–6. URL <https://arxiv.org/pdf/1208.4874.pdf>.
- [6] P. Fitzpatrick, Order conjugacy in finite groups, *Proc. Roy. Irish Acad. Sect. A* **85** (1985), 53–58.
- [7] T. M. Keller, Finite groups have even more conjugacy classes, *Israel J. Math.* **181** (2011), 433–444.
- [8] E. Landau, Über die Klassenzahl der binären quadratischen Formen von negativer Discriminante, *Math. Ann.* **56** (1903), 671–676. URL <https://eudml.org/doc/158080>.
- [9] G. A. Miller, Groups involving only a small number of sets of conjugate operators, *Arch. Math. und Phys.* **17** (1911), 199–204. URL <https://babel.hathitrust.org/cgi/pt?id=njp.32101033960566&view=1up&seq=221>.
- [10] J. Poland, Finite groups with a given number of conjugate classes, *Canadian J. Math.* **20** (1968), 456–464.
- [11] L. Pyber, Finite groups have many conjugacy classes, *J. London Math. Soc.* **46** (1992), 239–249.
- [12] A. Vera-López and J. Sangroniz, The finite groups with thirteen and fourteen conjugacy classes, *Math. Nachr.* **280** (2007), 676–694.

<sup>5</sup>For  $p = 2$  this group is abelian, since all groups of order 4 are, but for  $p \geq 3$  this group is nonabelian since  $f(x) = x + x^2$  and  $g(x) = x + x^3$  don't commute:  $f(g(x)) = x + x^2 + x^3 + 2x^4 \bmod x^5$  and  $g(f(x)) = x + x^2 + x^3 + 3x^4 \bmod x^5$ .